

Amélioration de la récupération du watermark par technique de l'étalement du spectre à fréquence porteuse variable

Tiberiu MUNTEAN, Eric GRIVEL, Mohamed NAJIM

Equipe Signal & Image, ENSEIRB, Université Bordeaux I

BP 99, F-33 402 Talence Cedex

tibi@tsi.u-bordeaux.fr

Résumé - Le tatouage de signaux audio est une technique qui permet de signer des media numériques, pour éventuellement protéger les droits d'auteurs, en incorporant une information inaudible dans le signal original et robuste aux attaques et traitements tels que le filtrage, la compression, etc. Nous proposons dans cette communication une procédure d'insertion qui utilise une technique hybride d'étalement de spectre par séquences directes et saut de fréquence adaptatif. De plus une clé secrète est utilisée pour améliorer le taux de récupération de l'information binaire représentant la marque.

Abstract – Audio watermarking is a way to insert inaudible information into audio signals in order to protect the ownership right. The insertion must be robust against usual signal processing (filtering, compression, etc.) and against attacks making the protection ineffective. The method proposed here uses an hybrid spread spectrum technique, direct sequence and adaptive frequency hopping. It is intended to improve the method proposed by Garcia et al. [6]. A secret key, generated during watermarking process, enables us to improve the bit error rate during the retrieval of hidden information.

1. Introduction

Lors de la dernière décennie, la numérisation des signaux audio a connu un essor considérable et a apporté un grand nombre d'avantages et de facilités pour distribuer et utiliser l'information. Mais la copie, sans perte de qualité, de fichiers son à l'aide de graveurs grand public ainsi que la multiplication des transferts et des téléchargements de fichiers MP3 par le net ont des effets néfastes, notamment dans le cadre de la protection des droits d'auteur.

Pour pallier ce problème, des recherches sur l'insertion d'informations cachées dans les documents numériques audio sont menées, à notre connaissance, depuis 1994 [2]. C'est la naissance du tatouage de documents, plus communément appelé *insertion de filigranes* ou encore *watermarking*.

Ces procédés doivent, tous, satisfaire certaines exigences notamment : (1) signal original doit subir des dégradations imperceptibles à l'utilisateur ; (2) la marque doit être statistiquement invisible ; (3) la marque doit être robuste aux traitements des signaux tels la compression, le filtrage, le codage/décodage, la requantification, etc. ainsi qu'aux tentatives intentionnelles de destruction ; (4) le procédé doit autoriser l'insertion de plusieurs watermarks différents pour permettre la distribution de média ; (5) l'introduction de la marque peut intervenir soit dans le signal original soit dans sa variante comprimée. Nous nous intéressons dans la suite à la première catégorie.

De nombreuses techniques ont été développées ces dernières années. Dans [1], Bassia et al. proposent un

algorithme qui modifie les caractéristiques temporelles du signal, en changeant le bit le moins significatif des données à marquer. Bien que cette technique soit considérée comme robuste devant une compression de type MPEG 2, son inconvénient majeur est sa faible résistance devant la requantification. Dans [2], les auteurs signalent que l'on peut exploiter la propriété que le système auditif humain (SAH) ne perçoit que la phase relative du signal.

Les techniques d'écho peuvent être également utilisées pour dissimuler une signature. Le codage s'effectue alors en créant un écho du signal original et en ajustant l'amplitude pour qu'il soit inaudible [8]. Cette approche a été reprise dans [11].

Dans [3] et [13], les auteurs proposent une insertion temporelle du watermark. Ce dernier est généré en filtrant une séquence pseudo-aléatoire par le masque fréquentiel du signal original. Le signal tatoué obtenu est ensuite pondéré par l'enveloppe du signal original, pour représenter le masquage temporel. Dans [7], Gomes et al. proposent d'imposer une propriété particulière au watermark, exploitant ainsi l'idée présentée par Furon et al. [4]. Les auteurs rendent le watermark cyclostationnaire, à partir d'une modulation d'amplitude et insèrent le watermark dans le domaine temporel du signal original. Cependant, il est à noter que dans une grande partie des techniques existantes se pose un problème de synchronisation pour récupérer le watermark à la réception. Une solution relative aux attaques désynchronisantes consiste à ajouter des séquences binaires [10], dites d'apprentissage, supposées connues.

Une autre approche de tatouage a été mise en œuvre par Garcia dans [6]. Il substitue les composantes fréquentielles situées sous le masque fréquentiel par celles du watermark. Il propose ensuite d'utiliser les techniques d'étalement de spectre pour améliorer la robustesse du watermark, en effectuant une modulation à phase binaire BPSK (Binary Phase Shift Keying). Cependant, l'identification du watermark dans le signal transmis doit être améliorée.

Dans cet article, nous proposons une méthode de tatouage de signaux audio par substitution fréquentielle reposant sur les travaux présentés dans [3], [13] et [6] qui combinent le masquage fréquentiel et l'étalement de spectre. Notre contribution consiste à introduire une fréquence porteuse qui varie selon le spectre de la trame analysée et à réaliser le marquage uniquement sur des trames qui satisfont les critères proposés.

Cet article s'articule en 4 parties. Dans la partie 2, nous rappelons la notion d'étalement de spectre. Dans la partie 3, nous présentons des améliorations de l'algorithme proposé à l'origine dans [6]. Des résultats expérimentaux sont fournis dans la partie 4.

2. Etalement de spectre

Les techniques d'étalement de spectre ont pour objet de modifier la représentation de l'information à transmettre en l'étalant sur une plage de fréquence plus grande que la bande de fréquence d'origine. Assimilée à un bruit, elle est donc plus difficile à détecter qu'un signal localisé dans une bande de fréquence donnée. Ces techniques ont été surtout utilisées dans le cadre des communications radio pour éviter l'interception du signal transmis. Il existe plusieurs méthodes pour étaler le spectre d'un signal donné :

- (α) la multiplication du signal à transmettre par une séquence pseudo-aléatoire binaire de fréquence beaucoup plus grande. Le spectre du signal résultant est ainsi obtenu par la convolution des deux spectres initiaux.
- (β) la méthode par saut de fréquence : le signal à transmettre module une porteuse de fréquence variable en temps selon une séquence pseudo-aléatoire. Le spectre résultant est alors le spectre du signal à transmettre positionné autour de différentes porteuses.

Puisque le watermarking peut être vu comme un système de communication, il est avantageux d'utiliser les techniques d'étalement du spectre pour l'insertion.

Soient le watermark w_n représenté par une séquence binaire dont la durée d'un bit est T_w et la séquence pseudo-aléatoire, p_n , dont la durée d'un bit est T_p beaucoup plus faible que T_w . En multipliant w_n par p_n , la largeur de bande de l'information transmise passe de $2/T_w$ à $2/T_p$.

En utilisant la modulation BPSK dont la fréquence porteuse instantanée est notée f_i , le signal résultant est de la forme suivante :

$$W(t) = \sqrt{2P} \sin\left(2\pi f_i t + \frac{w_n p_n \pi}{2}\right) = w_n p_n \sqrt{2P} \cos(2\pi f_i t)$$

où P est la puissance véhiculée par la porteuse.

La séquence pseudo-aléatoire, p_n , et la fréquence porteuse f_i , sont alors nécessaires pour effectuer les étapes d'étalement/désétalement, qui nécessitent aussi une bonne synchronisation.

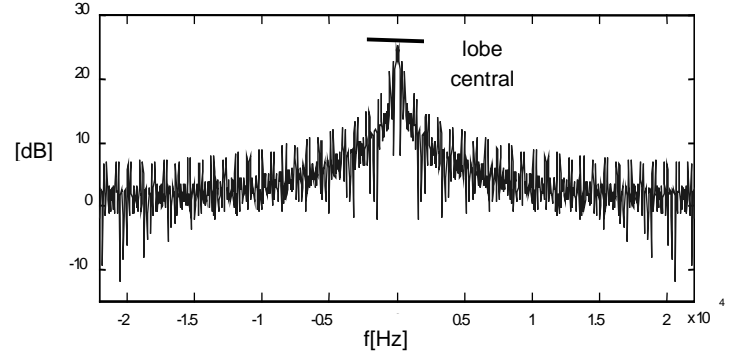


FIG. 1 : exemple du spectre étalé

La figure 1 met en évidence un lobe central pour le spectre étalé du watermark. Dans l'algorithme que nous proposons une technique hybride est utilisée, combinant les approches notées (α) et (β).

3. Deux critères pour améliorer l'insertion et la récupération du watermark

3.1 Présentation de l'algorithme

L'algorithme que nous proposons repose sur la substitution de composantes spectrales du signal original situées sous le masque fréquentiel $M(f)$ [5] par des composantes spectrales du watermark étalé [6] :

$$S_w(f) = \begin{cases} S(f) & \text{si } P_S(f) \geq M(f) \\ \alpha W(f) & \text{si } P_S(f) < M(f) \end{cases}$$

où $S_w(f)$ représente le spectre du signal marqué, $S(f)$ le spectre de la trame analysée du signal original, $W(f)$ le spectre du watermark étalé et $P_S(f)$ la densité spectrale du signal original.

Pour assurer une récupération efficace de la marque à la réception, il est cependant nécessaire que l'on dispose d'un certain nombre de ces composantes spectrales. Dans [6], la fréquence de la porteuse de la modulation BPSK est fixe, ce qui limite le nombre de composantes fréquentielles les plus significatives du watermark étalé inséré. Pour pallier ce problème, nous proposons d'adapter l'insertion du watermark selon les caractéristiques spectrales du signal original. Nous envisageons deux aspects :

- (1) le marquage n'est validé que pour des trames dans lesquelles on peut incorporer un nombre suffisamment grand de composantes spectrales du watermark par rapport à la largeur de bande de $W(t)$;
- (2) la fréquence de la porteuse f_i de la modulation BPSK varie et s'adapte selon la trame analysée. L'analyse du contenu fréquentiel du signal original fournit une

fréquence si la trame peut être marquée et une valeur préétablie dans les autres cas. Ces valeurs constituent alors une clé secrète de l'algorithme, dont la dimension est égale au nombre de trames traitées dans l'algorithme.

Il est à noter que l'utilisation d'une telle approche rend alors variable le débit d'insertion de l'information binaire représentant la marque.

3.2 Illustrations

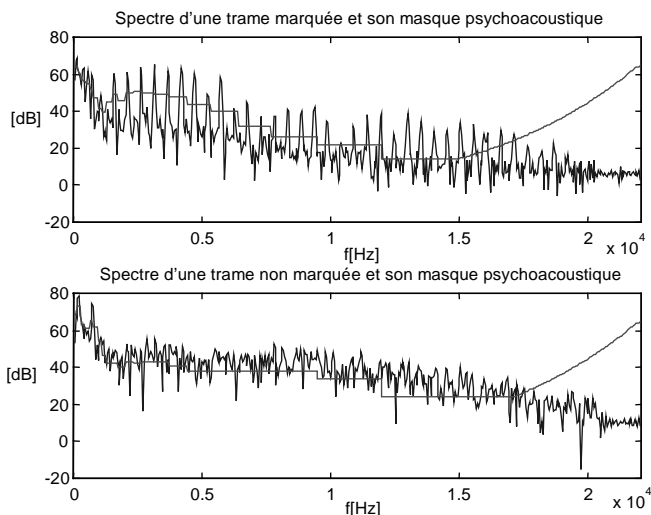


FIG. 2 : trame validée et trame invalidée

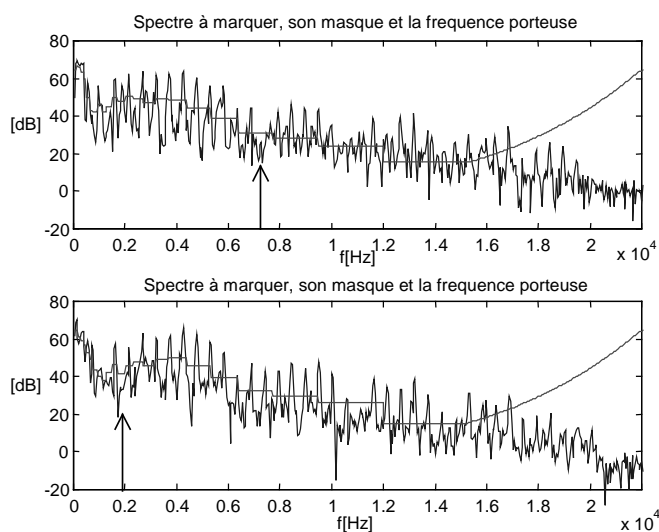


FIG. 3 : deux trames marquées et les fréquences des porteuses (représentées par des flèches)

Dès lors, deux critères pour choisir la fréquence de la porteuse sont envisagés :

- (C1) le premier consiste à trouver la bande de fréquence de largeur la plus grande possible et supérieure à un seuil prédéfini. La fréquence de la porteuse correspond alors à la fréquence centrale de cette bande.
- (C2) le second revient à étudier pour une largeur de bande de fréquence donnée, le rapport entre le nombre de composantes situées sous le masque et le nombre total

de composantes. On s'intéresse alors à la bande de fréquence pour laquelle ce rapport est le plus élevé.

3.3 Précision sur la génération de la séquence pseudo-aléatoire

La séquence pseudo-aléatoire utilisée pour étaler le spectre du watermark est obtenue en utilisant la clé secrète. L'algorithme utilisé est présenté à la figure 4 :

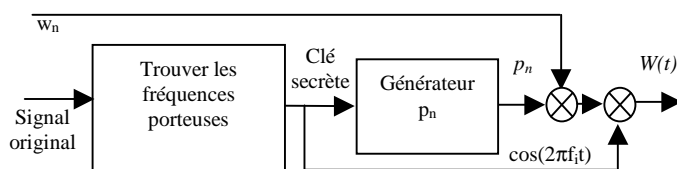


FIG. 4 : génération de la séquence p_n

Chaque valeur de la clé secrète est utilisée pour générer une suite pseudo-aléatoire de longueur égale à la celle de la trame d'analyse. A partir d'une concaténation de ces suites pseudo-aléatoires et d'une permutation aléatoire [9], nous obtenons la séquence pseudo-aléatoire utilisée pour étaler le spectre du watermark.

Avec cette procédure, la séquence pseudo-aléatoire peut être obtenue à partir de la clé secrète et de l'algorithme qui génère les suites aléatoires.

3.4 Récupération de la marque

La récupération de la marque ne nécessite pas la connaissance de la séquence pseudo-aléatoire. Seule la clé secrète permet de retrouver l'information cachée.

Les figures 5 et 6 présentent les schémas synoptiques de récupération et d'insertion de la marque.

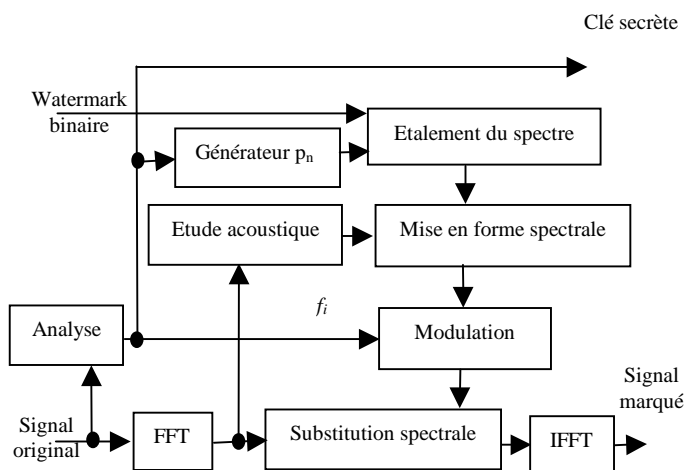


FIG. 5 : insertion du watermark

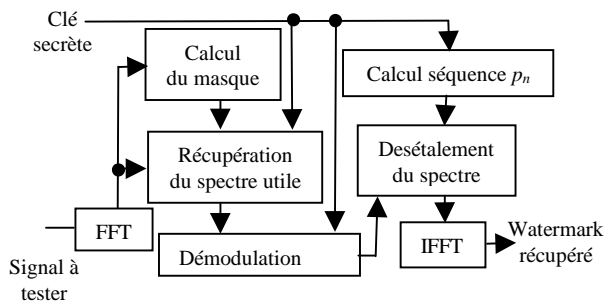


FIG. 6 : récupération du watermark

4. Résultats

Des tests ont été menés sur plusieurs signaux musicaux, échantillonnés à 44,1 kHz. L'information de watermark est représentée par une série de 50 valeurs binaires ayant la durée d'un bit de 10ms. La séquence pseudo-aléatoire binaire est utilisée à 1000b/s.

Quatre tests ont été réalisés :

- (1) tests acoustiques subjectifs pour vérifier l'inaudibilité des modifications apportées par le marquage ;
- (2) test pour étudier la capacité d'insertion dans différents types de signaux ;
- (3) test pour déterminer la relation qui existe entre le débit d'insertion de l'information binaire et le taux d'erreur des bits récupérés à la réception ;
- (4) test de robustesse à un bruit blanc additif et à un bruit coloré, pour simuler la robustesse vis à vis de la compression.

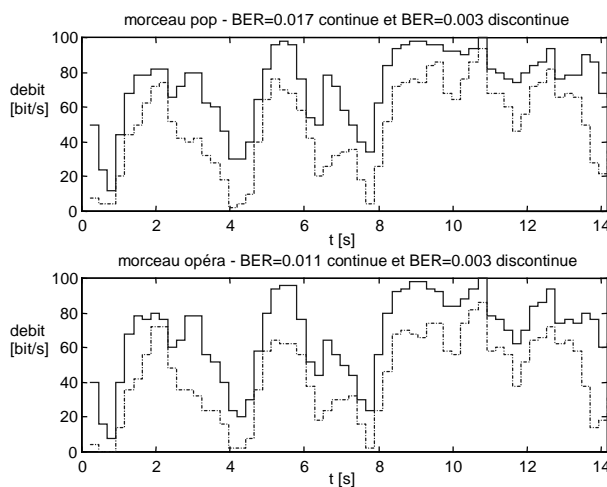


FIG. 7 : débit de l'information insérée

A la suite des tests subjectifs, nous avons constaté que la méthode d'insertion proposée ne modifie pas de façon perceptible le signal audio original. La capacité d'insertion du watermark varie avec le temps et reflète la nature non stationnaire du signal original.

A la figure 7, nous présentons le débit binaire en fonction du temps pour deux valeurs d'erreur de bit (BER) correspondantes. A la figure 8, le BER est présenté pour deux signaux différents et pour les deux critères (C1) et (C2) proposés dans le paragraphe 3.2.

Les techniques présentées ici fournissent un taux d'erreur plus faible que la méthode proposée dans [6].

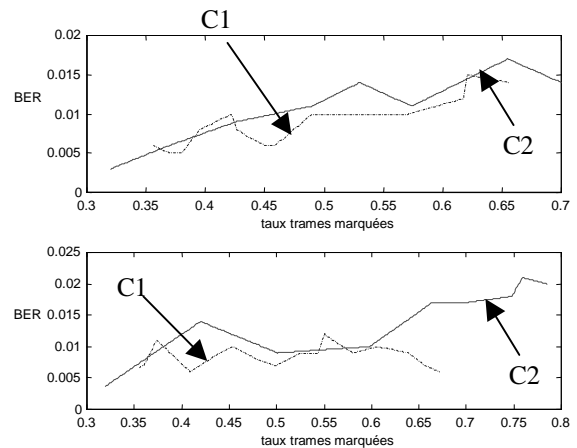


FIG. 8 : taux d'erreur de récupération binaire pour (a) morceau opéra, (b) morceau pop

La méthode présentée n'est pas sensible lorsque l'on ajoute un bruit blanc inaudible au signal marqué et, dans le cas du bruitage par bruit coloré, le BER est inférieur à 5%.

5. Références

- [1] P. Bassia, I. Pitas. *Robust Audio Watermarking in Time Domain*. EUSIPCO 1998, 8-11 Sept., Patras, Greece, pp. 25-28.
- [2] W. Bender, D. Gruhl, N. Morimoto. *Techniques for Data Hiding*, Technical Report, MIT Media Lab, 1994.
- [3] L. Boney, A. Tewfik, K. Hamdy. *Digital Watermarks for Audio Signals*. 1996 IEEE Int. Conf. on Multimedia Computing and Systems, Hiroshima, Japan 1996.
- [4] T. Furon, P. Duhamel, *An Asymmetric Public Detection Watermarking Technique*, Proc of the 3rd Int. Work of Information Hiding, Dresden, September 1999.
- [5] J. D. Johnston. *Transform Coding of Audio Signals Using Perceptual Noise Criteria*. IEEE Journal on Select. Areas in Comm., Vol. 6, No. 2, Feb. 1998.
- [6] R. A. Garcia, *Digital Watermarking of Audio Signals Using a Psycho-acoustic Auditory Model and Spread Spectrum Theory*, 107th AES Convention, Sept. 1999.
- [7] L. de C. T. Gomes, M. Mboup, M. Bonnet, N. Moreau, *Cyclostationarity-Based Audio Watermarking with Private and Public Hidden Data*, 109th AES Convention, Sept. 22-25, 2000.
- [8] D. Gruhl, A. Lu, W. Bender. *Echo Hiding*. Proceedings of the Workshop of Information Hiding, Cambridge, England, May 1996.
- [9] D. Knuth. *The Art of Computer Programming*. vol.II, Addison-Wesley, 1981.
- [10] N. Moreau, P. Dymarski, L. de C. T. Gomes. *Tatouage audio: Une réponse à une attaque désynchronisante*, CORESA 2000.
- [11] H. O. Oh, J. W. Seok, J. W. Hong, D. H. Youn. *New Echo Embedding Technique For Robust And Imperceptible Audio Watermarking*. Icassp 2001.
- [12] R. L. Pickholtz, D. L. Schilling, L. B. Milstein. *Theory of Spread Spectrum Communications-A Tutorial*, IEEE Transactions on Communications, May 1982.
- [13] M. D. Swanson, B. Zhu, A.H. Twefik, L. Boney, *Robust Audio Watermarking Using Perceptual Masking*, IEEE Signal Processing, Vol. n°66, pp. 337-355, 1998.